**SUPPLEMENTARY METHODS**

**Literature search**

Relevant literature cited throughout the manuscript were found with the aid of the Google Scholar search tool. Papers reflecting seagrass species’ response, taxon sampling, conservation and global change were identified using the Google Scholar (Accessed: August 2019). We used a search strategy that identified articles with “seagrasses”, “seagrass sampling”, “seagrass evolutionary diversity”, “seagrass conservation”, “seagrass and global change”, and “seagrass and extinction risk” as main subjects in the title, abstracts or keywords.

**Species geographic data**

# Seagrass occurrence data was downloaded from the Global Biodiversity Information Facility (GBIF; Data source: <https://doi.org/10.15468/dl.t7xgct> on January 17, 2020). A total of 72 seagrass species were included, resulting in a total of 938,000 occurrences. We also downloaded expert-based, range polygons for each species from the International Union for the Conservation of Nature (IUCN) Redlist website (<https://www.iucnredlist.org/resources/spatial-data-download>). Using these datasets, we quantify sampling effort to determine impediments that can limit understanding of seagrass species’ response to global change. All subsequent analyses were performed using R version 3.6.3 (R Core Team 2020).

**DATA ANALYSES**

**Assessment of uncertainties in geographic sampling**

Polygon shapefiles from the IUCN and point records from GBIF were overlaid onto 100 km × 100 km grid cells to assess coverage based on expert opinion versus actual sampling. We cleaned the data by removing extraneous occurrences that were inconsistent with known seagrass distributional range e.g., points falling on terrestrial areas, or those missing geographic coordinates or collection dates. We also standardized the species’ taxonomy against World Flora Online (www.worldfloraonline.org). Within each grid, the total amount of seagrasses sampled and the count for each individual species present was calculated. In the species richness plot, only the number of different species present within each grid was considered significant.

**Assessment of uncertainties in temporal sampling**

Seagrass collection dates from GBIF was converted into Julian Day of Year format (DOY; where 1 January = 1 DOY and 31 December = 365 DOY, etc.) using the R package “lubridate” (Grolemund & Wickham, 2011). We generated circular plots indicating the sampling density for any given month over the course of all years spanned by the dataset for each marine ecoregion. This approached allowed us to explore temporal sampling within each marine ecoregion as opposed to computing temporal sampling at a global scale. We then explored trends across decades by combining and plotting our datasets across centuries (1700-2000s) to assess overall changes in the density of seagrass sampling over decades.

**Assessment of uncertainties in taxonomic sampling**

We used the occurrence data from GBIF to quantify taxonomic/phylogenetic bias, the tendency of some species or clades to be collected more or less than others. If closely related species are likely to be sampled similarly than expected by chance, then phylogenetic bias results. Along these lines, we obtained a dated phylogenetic tree of seagrass from Daru et al. (2017) to quantify phylogenetic bias in the number of occurrences collected per species. We then assessed phylogenetic signal in the number of specimens collected per species to explore if closely related seagrass species were collected similarly than expected by chance. If phylogenetic signals were detected, this would provide evidence that the sampling of seagrass species were concentrated in specific lineages to the point that closely related species tended to be sampled similarly (Plazzi et al., 2010). Phylogenetic signal test was performed using Moran’s I, Abouheif’s C-mean, Pagel’s λ, and Blomberg’s K. The values of Pagel’s k range from 0 (no phylogenetic signal) to 1 (strong phylogenetic signal) and is a scaling parameter for the phylogeny and measures phylogenetic dependence of observed trait data (Pagel, 1999). For Blomberg’s K, a significant phylogenetic signal, is indicated by a K value > 1. These were calculated using the R package “phytools” (Revell, 2012).

**Rank correlations of threatened seagrass families across marine ecoregions**

We used a pairwise rank correlation analysis to calculate the correlation of seagrass families possessing threatened species across different marine ecoregions of the world (MEOWs). Correlational values were assigned based on the level of overlap of seagrass families with threatened species between MEOWs. Low correlation values between MEOWs indicate that the threatened seagrass species in these regions are unique to those ranges, while larger correlational values indicate MEOWs share a significant amount of seagrass families possessing threatened species.

**Assessment of taxonomic distribution of extinction risk**

Each seagrass species was grouped in the following threat categories: Least Concern (LC), Lower Risk/Conservation Dependent (LR/CD), Near Threatened (NT), Vulnerable (VU), Endangered (EN), Critically Endangered (CR), and Extinct (EX) (assessed from the IUCN August 2019). We further grouped these categories into two broader categories: threatened (EX + CR + EN + VU) and non-threatened (LR/CD + NT + LC). Data deficient species were removed from prior to analysis. We explored the proportion of threatened vs. non-threatened species within each seagrass family by dividing the number of threatened species in a family by the total number of species in that family. Significance was assessed by shuffling species membership among families 1000 times and recalculating the ratio of threatened species within each random draw while maintaining the number of species per family constant. The observed proportion of threatened species was compared to the random draws to obtain p values.

**DATA AVAILABILITY STATEMENT**

The raw data underlying most analyses in this paper (Figures 4, 5, 6, and 9) was derived from the global biodiversity information facility (GBIF: <https://doi.org/10.15468/dl.t7xgct>). IUCN data used for Figure 8 is available for download at: <https://www.iucnredlist.org/resources/spatial-data-download>. The source data underlying Figures 3 and 7 are provided as a Source Data files.

**CODE AVAILABILITY STATEMENT**

All scripts and code necessary to repeat the analyses described here have been made available in the following vignette: https://github.com/brirock35/Impediments-to-Understanding-Seagrasses-Response-to-Global-Change/blob/main/README.md.

**SUPPLEMENTARY REFERENCES**

Daru, B. H., Holt, B. G., Lessard, J.-P., Yessoufou, K., & Davies, T. J. (2017). Phylogenetic regionalization of marine plants reveals close evolutionary affinities among disjunct temperate assemblages. *Biological Conservation*, *213*, 351–356. https://doi.org/https://doi.org/10.1016/j.biocon.2016.08.022

Grolemund, G., & Wickham, H. (2011). Dates and times made easy with lubridate. *Journal of statistical software*, *40*, 1-25. http://hdl.handle.net/10.18637/jss.v040.i03

Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature, 401,* 877-884. https://doi.org/10.1038/44766

Plazzi, F., Ferrucci, R. R., & Passamonti, M. (2010). Phylogenetic representativeness: a new method for evaluating taxon sampling in evolutionary studies. *BMC Bioinformatics*, *11*, 209. https://doi.org/10.1186/1471-2105-11-209

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/

Revell L. J (2012). “phytools: An R package for phylogenetic comparative biology (and other things).” Methods in Ecology and Evolution, *3*, 217-223. https://doi.org/10.1111/j.2041-210X.2011.00169.x